

SOEPpapers

on Multidisciplinary Panel Data Research

303

Christian Pihl • Ulrich Pötter

**Möglichkeiten und Grenzen der partiellen Rangdaten-
analyse in der sozialwissenschaftlichen Forschung
– Untersuchungen zu sozialen Netzwerken und Hilfeleistungen –**

Berlin, Mai 2010

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
<http://www.diw.de/soeppapers>

Editors:

Georg **Meran** (Dean DIW Graduate Center)

Gert G. **Wagner** (Social Sciences)

Joachim R. **Frick** (Empirical Economics)

Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics)

Christoph **Breuer** (Sport Science, DIW Research Professor)

Anita I. **Drever** (Geography)

Elke **Holst** (Gender Studies)

Martin **Kroh** (Political Science and Survey Methodology)

Frieder R. **Lang** (Psychology, DIW Research Professor)

Jörg-Peter **Schräpler** (Survey Methodology)

C. Katharina **Spieß** (Educational Science)

Martin **Spieß** (Survey Methodology, DIW Research Professor)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

Möglichkeiten und Grenzen
der partiellen Rangdatenanalyse in der
sozialwissenschaftlichen Forschung
– Untersuchungen zu sozialen Netzwerken und Hilfeleistungen –

von Christian Pihl¹ und Ulrich Pötter²

¹ *Westfälische Hochschule Zwickau, Fachbereich Gesundheits- und Pflegewissenschaften*

² *DJI – Deutsches Jugendinstitut München, Zentrum für Dauerbeobachtung und Methoden*

Abstract

Pflegebedürfnisse werden zum erheblichen Teil durch professionelle Dienste befriedigt. Ein wichtiger Eckpfeiler in der pflegerischen Versorgung stellen darüber hinaus familiäre Beziehungen wie Partner und Kinder, sowie Freundes- und Nachbarschaftsnetzwerke dar. Diese Ressourcen stehen aber nicht allen Personen gleichermaßen zur Verfügung, sondern sind in der Bevölkerung ungleich verteilt.

Wir beschäftigen uns mit der Reihenfolge von genannten Unterstützungspersonen, bei der Befragte Personengruppen – überwiegend aus dem Kreis informeller Beziehungen – im Hinblick auf die Versorgung im Rahmen einer hypothetischen Pflegebedürftigkeit nennen sollten. Es handelt sich um eine Frage, die u.a. den Befragten des Sozio-ökonomischen Panels regelmäßig gestellt wird. Die Besonderheit der Situation einer pflegerischen Bedürftigkeit aufgrund der Notlage und der Intimität des Ereignisses setzt eher vertrauensvolle Personen als Ansprechpartner voraus. Inhaltlich können somit durch diese Personenangaben Hinweise auf die Ausgestaltung individuell differenzierter Teilnetzwerke gefunden werden.

Ähnliche Fragen nach Präferenzen finden sich aber auch in vielen anderen Bereichen der Sozialstatistik, in denen Befragte gebeten werden, ihre Präferenzen für bestimmte Alternativen zu äußern. Die besonderen Eigenschaften solcher Daten erfordern geeignete Methoden, die den Besonderheiten der erzeugten Daten Rechnung tragen. Die vorliegende Arbeit soll einen Beitrag zur Darstellung und Analyse solcher Rangdaten leisten. Dabei werden zwei Ziele verfolgt: Einerseits wird ein Verfahren zur Visualisierung von Rangdaten dargestellt, das bislang in der sozialwissenschaftlichen Forschung wenig Beachtung gefunden hat und hier im konkreten Fall für die Darstellung von Teilnetzwerken angewendet wird. Es handelt sich um die Darstellung von Rangdaten mit Hilfe von Polytopen, durch die sich die Komplexität in den Präferenzäußerungen explorativ grafisch veranschaulichen lässt. Andererseits soll anhand dieses Verfahrens die inhaltliche und methodische Problematik einer Fragebogenkonstruktion diskutiert werden, in der Personen gebeten werden, Aussagen über hypothetische Zustände oder Meinungen zu leisten, um ihren Präferenzen Ausdruck zu verleihen.

1 Einleitung

Wir beschäftigen uns mit der Reihenfolge, in der Befragte Personengruppen in einer Notlage um Unterstützung bitten. Es handelt sich um eine Frage, die u.a. den Befragten des Sozio-ökonomischen Panels regelmäßig gestellt wird. Ähnliche Fragen nach Präferenzen finden sich aber auch in vielen anderen Bereichen der Sozialstatistik, in denen Befragte gebeten werden, ihre Präferenzen für bestimmte Alternativen zu äußern. Die besonderen Eigenschaften solcher Daten erfordern geeignete Methoden, die den Besonderheiten der erzeugten Daten Rechnung tragen. Die vorliegende Arbeit soll einen Beitrag zur Darstellung und Analyse solcher Rangdaten leisten. Dabei werden zwei Ziele verfolgt: Einerseits wird ein Verfahren zur Visualisierung von Rangdaten dargestellt, das bislang in der sozialwissenschaftlichen Forschung wenig Anwendung gefunden hat. Es handelt sich um die Darstellung von Rangdaten mit Hilfe von Polytopen, durch die sich die Komplexität in den Präferenzäußerungen explorativ veranschaulichen lässt. Andererseits soll anhand dieses Verfahrens die inhaltlich und methodische Problematik einer Fragebogenkonstruktion diskutiert werden, in der Personen gebeten werden, Aussagen über hypothetische Zustände oder Meinungen zu leisten, um ihren Präferenzen Ausdruck zu verleihen.

Die Arbeit gliedert sich wie folgt: Zunächst werden im folgenden Kapitel Besonderheiten von Rangdaten dargestellt. Methoden und Datenbeschreibung der empirischen Analyse finden sich in Kapitel 3. Die deskriptive Beschreibung von ausgewählten Rangdaten im Rahmen einer explorativen Darstellung erfolgt in Kapitel 4. Die Arbeit schließt mit einer Zusammenfassung in Kapitel 5.

2 Besonderheiten von Rangdaten

Die Frage nach bevorzugten Personengruppen, an die man sich im Fall einer Notlage wenden würde und die Angabe einer entsprechenden Rangfolge zwischen den Personengruppen setzt zunächst eine Einteilung der unterstützenden Personen in Kategorien, etwa in Partner, Verwandte, Freunde, professionelle Helfer etc. voraus. Nur eine solche Abstraktion von je konkreten Personen im Umfeld des Befragten erlaubt die Konstruktion von Präferenzordnungen, die für die Untersuchung sozialer Netzwerke relevant sein können. Sind etwa als mögliche Personengruppen die vier Kategorien Partner, Verwandte, Freunde und professionelle Helfer vorgegeben, kann man die Befragten bit-

ten, sie in eine Reihenfolge zu bringen: Welche Personengruppe würde der Befragte als erste um Unterstützung bitten, an wen würde er sich dann wenden etc. Man erhält eine Ordnung der Personengruppen, etwa (Partner, Verwandte, Freunde, professionelle Helfer) oder (Partner, Freunde, Verwandte, professionelle Helfer). Schreibt man allgemein $\{O_1, O_2, \dots, O_m\}$ für die verschiedenen Personengruppen (oder eine beliebige Menge von m Objekten), dann ist der *Ordnungsvektor* der Vektor der Indizes der Personengruppen in der Reihenfolge der Präferenz. Vergibt man etwa den Gruppen (Partner, Verwandte, Freunde, professionelle Helfer) die Nummern (1,2,3,4), dann ist der Ordnungsvektor der Reihenfolge (Partner, Verwandte, Freunde, professionelle Helfer) gerade (1,2,3,4), dagegen der der Reihenfolge (Partner, Freunde, Verwandte, professionelle Helfer) (1,3,2,4).

Die Ordnung lässt sich aber auch vollständig durch den *Rangvektor* wiedergeben: Der Präferenzordnung (Partner, Freunde, Verwandte, professionelle Helfer) entspricht der Rangvektor (1,3,2,4), weil etwa die Gruppe der Verwandten (an der zweiten Stelle der ursprünglichen Aufzählung) den dritten Rang in der Präferenzordnung belegen. Dies ist in diesem Beispiel identisch mit dem entsprechenden Ordnungsvektor. Für die Präferenzordnung (Partner, Freunde, professionelle Helfer, Verwandte) ergibt sich dagegen ein Rangvektor von (1,4,2,3), weil z.B. an der zweiten Stelle der Rang der Verwandten, also 4, angegeben wird. Der Ordnungsvektor ist aber (1,3,4,2), weil der Index der zweiten Präferenz (Freunde) gerade 3 ist. Um beide Formen der Angabe zu unterscheiden, werden wir im folgenden $(1,4,2,3)_R$ für Rangvektoren bzw. $(1,3,4,2)_O$ für die Ordnungsvektoren schreiben.

Hat man für eine (Teil-) Menge von Befragten ihre Präferenzordnungen erfragt (und entweder als Rang- oder Ordnungsvektoren kodiert), dann kann der resultierende Datensatz mit allen Techniken der multivariaten Statistik beschrieben werden. Insbesondere könnte man etwa die Häufigkeiten der ersten Nennungen, dann die der zweiten etc. schrittweise analysieren. Ebenso kann man die (durchschnittliche) Anzahl von Alternativen i angeben, die der Alternativen k vorgezogen werden. Dies ergibt eine $m \times m$ Matrix. Ähnlich kann man auch eine entsprechende Matrix berechnen, in der eingetragen wird, wieviele Personen Alternative i den Rang j zuweisen. Aber alle diese Möglichkeiten beschränken sich auf Teilaspekte der Präferenzordnung, indem sie nur einzelne Ränge betrachten, paarweise Alternativen vergleichen oder die Verteilung von Rängen je einer Alternativen angeben. Solche Verfahren sind für einen ersten Überblick über die Daten sehr wichtig. Da sie sich auf Teilaspekte der Präferenzordnungen

beschränken und notwendigerweise die Werkzeuge der klassischen multivariaten Statistik benutzen, werden einige offensichtliche Eigenschaften von Präferenzdaten ignoriert. Wir werden ein einfaches graphisches Verfahren vorstellen, dass es erlaubt, fast alle Eigenschaften der Verteilung von Präferenzdaten wiederzugeben.

3 Methoden und Datenbeschreibung

Als Datengrundlage für die vorliegende empirische Untersuchung dient das Sozioökonomische Panel (SOEP)¹, das am DIW Berlin seit 1984 jährlich als repräsentative Längsschnittuntersuchung im Paneldesign durchgeführt wird. Befragt werden alle Personen ab dem 16ten Lebensjahr, die Mitglieder eines Haushaltes in der Stichprobe sind. Mittels eines Personenfragebogens werden umfangreiche individuelle Merkmale und Einschätzungen erfasst. Zusätzlich ist vom Haushaltsvorstand ein Haushaltsfragebogen auszufüllen, mit dem der jeweilige Haushaltskontext und haushaltsbezogene Informationen erhoben werden.

Die hier bei der Datenanalyse verwendete Stichprobe bezieht ihre Daten aus einer Personenbefragung des SOEP aus dem Jahr 2006. Obwohl die Fragestellung auch in den Jahren 1991, 1996 und 2001 verwendet wurde, beschränken wir die Querschnittsanalyse auf das Jahr 2006. Dies bietet den Vorteil, dass alle Ergänzungsstichproben berücksichtigt werden können, die im Laufe der Jahre ergänzend gezogen wurden. Die Angaben stammen von Personen der Teilstichproben A (Teilstichprobe West), B (Teilstichprobe Ausländer) und C (Teilstichprobe Ost, ab 1990), D (Zuwandererstichprobe), E (Erneuerungsstichprobe, 2006) und F (Hocheinkommensstichprobe).

3.1 Die abhängige Variable

Ausgangspunkt der folgenden inhaltlichen Ausführungen bilden Überlegungen zur Sozialkapitalmessung, die im SOEP implementiert wurden (vgl. Diwald et al. 2006; Sattler und Diwald 2009). Unter dem sozialen Kapital eines Individuums können die Erträge bzw. Ressourcen verstanden werden, die sich aus zwischenmenschlichen Beziehungen ergeben. Die Gesamtheit des Ressourcenpotentials hängt direkt vom sozialen Netzwerk jedes Individuums ab. Allerdings lässt sich die Gesamtheit der vielfältigen Ertragsmöglichkeiten des Sozialkapitals aus den Beziehungen in einem sozialen Netzwerk

¹ Eine detaillierte Beschreibung des SOEP findet sich in: Wagner et al. (2008).

nur schwer erfassen, so dass es sich anbietet, den Fokus auf einen Teilertrag zu legen.

3.2 Die inhaltliche Dimension

Einen solchen Teilertrag des sozialen Kapitals bilden Unterstützungsleistungen im Falle einer längerfristigen Pflegebedürftigkeit. Die Variable, die in dieser Arbeit Gegenstand der Untersuchung ist, bezieht sich auf das von den Befragten angegebene Unterstützungspotential für den Fall eines solchen Bedarfs. Die Frage, die den teilnehmenden Personen gestellt wurde, ist folgendermaßen eingeleitet:

„Auf dieser Liste stehen Personen, die für Sie in irgendeiner Weise bedeutsam sein können. Wie ist es bei Ihnen, wenn es um die folgenden Dinge geht? Nennen Sie bitte jeweils bis zu drei Personen von der Liste.“ (Personenfragebogen 2006, Frage 115)

Eine der darauf folgenden Fragen lautet²:

„Nur einmal hypothetisch gefragt: Wie wäre es bei einer langfristigen Pflegebedürftigkeit, z. B. nach einem schweren Unfall: Wen würden Sie um Hilfe bitten?“ (Personenfragebogen 2006, Frage 115c)

Die hypothetische Fragestellung bezieht sich auf eine konstruierte Situation, die bei den Befragten real nicht vorliegen oder bisher nicht aufgetreten sein muss. Damit müssen auch keine eigenen Erfahrungswerte in Bezug auf eine solche Situation existieren. Es handelt sich somit nicht um einen konkreten Ertrag, sondern um einen Erwartungswert.

Dies kann folgendermaßen interpretiert werden: In der Vergangenheit wurde bereits durch eine Investitionstätigkeit in Beziehungen Sozialkapital aufgebaut (vgl. umfassend: Pihl 2008). Dabei wurde ein soziales Netzwerk bestehend aus unterschiedlichen Beziehungen mit mehr oder minder großen Verpflichtungen gebildet. Der Befragte muss nun vor dem Hintergrund seiner Situation zum Zeitpunkt der Befragung eine Bewertung vornehmen und überprüfen, zu welchen Personen bislang eine Beziehung aufgebaut wurde, die qualitativ und quantitativ zur Unterstützung der Versorgung beitragen könnte. Existieren mehrere Möglichkeiten einer alternativen Nennung, muss eine Rangfolge der Personen festgelegt werden. Die Unterstützung bei einem pflegeri-

2 Weitere Fragen in diesem Kontext sind: Frage 115 a): „Mit wem teilen Sie persönliche Gedanken und Gefühle oder sprechen über Dinge, die Sie nicht jedem erzählen würden?“ Frage 117 b): „Wer unterstützt Sie in Ihrem beruflichen Fortkommen oder ihrer Ausbildung und hilft Ihnen, dass Sie vorankommen?“ Frage 117 d): „Wer kann Ihnen auch mal unangenehme Wahrheiten sagen?“

schen Bedarf als besondere Ressource bildet einen möglichen Teilertrag des Sozialkapitals und setzt eine entsprechende Bildung von Beziehungen zu einem früheren Zeitpunkt voraus.

Da es sich um eine Analyse von Erträgen auf der Erwartungsebene handelt, kann inhaltlich nicht darauf geschlossen werden, welche konkreten Unterstützungsleistungen von den Befragten gefordert werden. Tatsächlich kann sich hierunter ein höchst heterogenes Leistungsspektrum verbergen. Bei dieser Form einer schwereren körperlichen Einschränkung sind die benötigten Unterstützungsleistungen zeitintensiv, von intimer Natur und erfordern spezielle Fähigkeiten, wenn die pflegerische Versorgung von den Befragten als Unterstützung erwartet wird. Der zur Verfügung stehende Personenkreis für diese Tätigkeiten dürfte sich dann auf einen relativ kleinen Personenkreis beschränken und sehr verlässliche und stabile soziale Beziehungen voraussetzen. Da sich die Pflegeabhängigkeit in der Regel über einen längeren Zeithorizont erstreckt, ist der zu erbringende Aufwand für den Pflegenden mit hohen direkten Kosten und Opportunitätskosten verbunden.

Allerdings können aus Sicht der Befragten auch andere Dimensionen des Ertrages mit der Nennung verbunden sein. Durchaus denkbar ist hier beispielsweise, dass die eigentliche pflegerische Versorgung über Marktsubstitute erworben wird und der genannten Person die Aufgabe zugedacht wird, das Management der Pflege zu übernehmen. Der jeweilige Partner müsste sich dann z. B. um die Informationsbeschaffung, die Auswahl eines geeigneten Pflegedienstes und die Kontrolle der Versorgung kümmern. Da Pflege ein Vertrauensgut ist, kann mittels einer solchen Vorgehensweise über eine nahestehende Person, zu der es eine solche Vertrauensbasis gibt, eine qualitativ und quantitativ höchstmögliche Versorgung sichergestellt werden.

Bei den hier geschilderten unterschiedlichen Erwartungen hinsichtlich der einzubringenden Ressourcen des Anderen wird dennoch deutlich, dass die Nennung einer Person an einen bestimmten Beziehungstyp geknüpft ist. Angesichts des Vertrauensgutaspektes muss davon ausgegangen werden, dass nur Personen genannt werden, zu denen eine Beziehung vorhanden ist, deren zentrale Eigenschaft eine notwendige gemeinsame Vertrauensbasis ist. Tendenziell dürften somit eher langfristige und stabile Beziehungspartner angegeben werden.

Aus der Nennung einer Person als Ansprechpartner kann nicht darauf geschlossen werden, ob diese auch tatsächlich als Unterstützungspartner fungieren würde. Diewald un-

terscheidet deshalb zwischen einer strukturellen Perspektive, die letztlich die quantitative Dimension sozialer Netzwerke darstellt, und einer funktionalen Perspektive, in der die tatsächlich beobachtbaren Inhalte und geleisteten Transaktionen berücksichtigt werden (Diewald 1991, S. 78). Ebenfalls liegen keine Informationen darüber vor, ob die Genannten zeitlich, räumlich sowie hinsichtlich ihrer Fähigkeiten den Anforderungen einer pflegerischen Tätigkeit gewachsen sind, wie es von den Auskunft gebenden erwartet wird. In Anlehnung an die von Cohen und Syme (1984, S. 15) vorgenommene Unterscheidung verschiedener Merkmalsdimensionen sozialer Unterstützung kann somit zusammenfassend festgestellt werden, dass bei einer Nennung zwar eine subjektive Erwartung bezüglich einer Unterstützung vorliegt und auch eine Vorstellung darüber besteht, welche Leistungen damit verbunden sind. Dies bedeutet jedoch nicht notwendigerweise, dass diese Potentiale gleichermaßen auch objektiv vorhanden sind. Das sollte in der Interpretation der Ergebnisse Berücksichtigung finden.

3.3 Die Antwortkategorien

Bei der hier verwendeten Fragestellung gibt es drei Antwortmöglichkeiten, mittels derer dezidiert alternative Personen genannt werden konnten, die bedeutsame Personen des eigenen Lebens darstellen und in der spezifischen Lebenssituation mit dem Bedarf von Unterstützung bei einer pflegerischen Abhängigkeit als primäre Ansprechpartner dienen. In Tabelle 3-1 werden die gegebenen Antwortkategorien und die Häufigkeiten ihrer Nennungen dargestellt. Der Personenkreis wird auf die Untersuchungspersonen in der zweiten Lebenshälfte ab dem 40sten Lebensjahr eingeschränkt, wobei davon ausgegangen wird, dass bei dieser Altersgruppe bereits soziale Netzwerke im Hinblick auf die Fragestellung aufgebaut wurden.

Tabelle 3-1: Kodierung der abhängigen Variable

Originalvariable	1. An- sprech- person	%	2. An- sprech- person	%	3. An- sprech- person	%	Generiert
Ehepartner/ Partner	10068	73,0	348	4,0	107	2,5	(Ehe-) Partner
ehemaliger (Ehe-)Partner	71	0,5	36	0,4	11	0,3	
Mutter	394	2,9	900	10,4	164	3,8	Eltern
Vater	25	0,2	171	2,0	241	5,6	
Stiefvater, Pflegevater	9	0,1	5	0,1	3	0,1	
Stiefmutter, Pflegermutter	2	0,0	9	0,1	5	0,1	
Schwiegermutter	11	0,1	83	1,0	74	1,7	
Schwiegervater	5	0,0	17	0,2	33	0,8	
Sohn	514	3,7	1752	20,3	1310	30,2	Kind(er)
Tochter	1388	9,7	2960	34,3	462	10,7	
Bruder	70	0,5	192	2,2	252	5,8	Sonstige Verwandte
Schwester	255	1,9	485	5,6	384	8,9	
Enkel	23	0,2	64	0,7	95	2,2	
Großmutter	0	0,0	2	0,0	1	0,0	
Großvater	0	0,0	1	0,0	2	0,1	
Tante, Nichte	36	0,3	62	0,7	41	1,0	
Onkel, Nefte	12	0,1	19	0,2	18	0,4	
andere weibl. Verwandte	64	0,5	163	1,9	161	3,7	
andere männl. Verwandte	12	0,1	44	0,5	65	1,5	
Arbeitskollegen	15	0,1	26	0,3	28	0,7	Freunde, Nachbarn und Sonstige
Vorgesetzte	1	0,0	5	0,1	2	0,1	
Personen aus der Ausbil- dung	5	0,0	9	0,1	12	0,3	
Nachbarn	64	0,5	121	1,4	123	2,8	
Personen aus Verein, Freizeit	16	0,1	37	0,4	64	1,5	
Haushaltshilfe, Sozialhel- fer, ambulanter Dienst, bezahlte Helfer, u. ä.	802	5,8	1121	13,0	674	15,6	Substitut
Niemand	905			6,1			Niemand
n	14767		8632		4332		

Quelle: SOEP 2006, ungewichtete Daten, über 40-jährige Personen, Absolut- und Prozentwerte.

Gemessen an der Anzahl der Antwortkategorien liegt der Schwerpunkt auf möglichen Nennungen von Verwandten als Ansprechpartner, wobei der Verwandtschaftsgrad genau erfasst wird. Eine gewisse Inkonsistenz kann auf den ersten Blick in den möglichen Antworten gesehen werden, da nach den Schwiegereltern, nicht aber nach der Schwiegertochter und dem -sohn gefragt wird und diese unter den sonstigen Verwandten subsumiert werden. Abgesehen vom verwandtschaftlichen Kreis werden über fünf Items noch Beziehungspartner aus verschiedenen freundschaftlichen Netzwerken abgefragt, die jeweils unterschiedlichen Kontexten wie z. B. Beruf oder Nachbarschaft entstammen. Eine weitere Möglichkeit besteht in der Angabe von bezahlten Diensten als Ansprechpersonen. Zusätzlich kann „Niemand“ angegeben werden, wenn weder für die erste, noch für die zweite oder dritte Nennung ein Ansprechpartner vorhanden ist³.

In der letzten Spalte (Generiert) von Tabelle 3-1 ist die Kodierung dargestellt, wie sie in dieser Arbeit Anwendung findet. Die zur Verfügung stehenden, stark differenzierenden Items wurden zu sechs homogenen Kategorien zusammengefasst. Aufgrund der geringen Fallzahlen in einigen Kategorien bietet sich die Zusammenfassung zu Gruppen an, auch wenn damit Informationsverluste in Kauf genommen werden müssen. Um eine weiterführende Analyse zu gewährleisten, die den individuellen Unterschieden in der Sozialkapitalausstattung Rechnung trägt, stellt dies eine notwendige Einschränkung dar. Im Rahmen dieser Analyse werden somit die Informationen verdichtet und Nennungen von (Ehe-) Partnern, Kindern, Eltern, Verwandten, Freunden/Nachbarn und Sonstigen untersucht.

Ein Kennzeichen der bisher beschriebenen Personengruppen ist die Zugehörigkeit zum informellen Unterstützungspotential, das von der formellen pflegerischen Versorgung im Hinblick auf konkrete Leistungen dahingehend abgegrenzt werden kann, dass es sich um Tätigkeiten handelt, „für die (a) kein Lohn gezahlt wird und/oder (b) für die keine Steuern oder Sozialversicherungsbeträge entrichtet werden“ (Hank et al. 2008, S. 11). Inhaltlich hiervon abweichend sind zwei weitere Antworten genauer zu untersuchen:

Eine erste Besonderheit besteht in der Antwortmöglichkeit „Niemand“, da dieses Item nur genutzt werden kann, wenn über alle drei Nennungen weder eine Person des informellen Netzwerkes noch ein bezahlter Helfer im Sinne der Fragestellung als bedeutsam empfunden wurde.

3 Da es sich bei dieser Antwortmöglichkeit nicht um eine Rangordnung handelt, findet sie in der grafischen Darstellung keine Berücksichtigung, muss hier aber in der inhaltlichen Diskussion thematisiert werden.

Die stark sinkenden Fallzahlen (vgl. Tabelle 3-1) über die drei Antwortkategorien lassen vermuten, dass diese Form der Antwortmöglichkeit bei der zweiten und dritten Person die Anzahl der Befragten, die keinen weiteren Ansprechpartner nennen, unangemessen vernachlässigt. Verfügen Befragte lediglich über eine oder zwei Ansprechpartner und wollen aus persönlichen Gründen keine marktfähige Substitutionsmöglichkeit angeben, existiert keine weitere Möglichkeit der Nennung, die dies deutlich macht. Da sich die Fallzahlen deutlich verringern, sind Ausfälle auch wegen der fehlenden Aufsplittung der Antwortmöglichkeit „Niemand“ zu vermuten. Falls die Befragten nur eine Person nennen, kann nicht zweifelsfrei geschlossen werden, dass keine weiteren Unterstützungspersonen mehr vorhanden sind. Grundsätzlich ist ein Ignorieren der weiteren Fragestellung denkbar, da es sich um eine rein hypothetische Fragestellung handelt. Wenn sich beispielsweise der Befragte mit der inhaltlichen Dimension der Fragestellung nicht auseinandergesetzt hat, erscheint es durchaus plausibel, dass er eine erste Ansprechperson nennt, dann aber mit der Beantwortung überfordert ist und sie aus diesem Grunde verweigert.

Eine zweite Gruppe möglicher Antworten ist in der Nennung „Haushaltshilfe, Sozialhelfer, ambulanter Dienst, bezahlte Helfer, u. ä.“ (nachfolgend als Substitut bezeichnet) zu sehen, da es sich hier um ein Substitut bzw. marktfähiges Äquivalent der informellen pflegerischen Versorgung handelt. Eine solche Angabe muss nicht zwangsläufig bedeuten, dass dies aus eigenen Mitteln bezahlt wird. Eine Haushaltshilfe kann bei Vorliegen gewisser Voraussetzungen auch über die gesetzliche Krankenversicherung (GKV) gestellt oder erstattet werden.⁴ Im Sinne der Fragestellung, in der eine Situation nach einem schwereren Unfall beschrieben wird, kann eine vorausgehende stationäre Krankenhausbehandlung angenommen werden. Ein Teil der Befragten wäre also anspruchsberechtigt. Ob diese Leistung der GKV den Befragten bekannt ist und inwieweit dies bei der Nennung eines Substituts eine Rolle spielt, kann nicht ermittelt werden.

Inhaltlich bieten sich verschiedene Interpretationsmöglichkeiten hinsichtlich der Antworten „Niemand“ und eines Substituts an:

- Es wurden keine Investitionen in Beziehungen getätigt, mittels derer die hier

4 §38 SGB V: (1) Versicherte erhalten Haushaltshilfe, wenn ihnen wegen Krankenhausbehandlung oder wegen einer Leistung nach §23 Abs. 2 oder 4, §§24, 37, 40 oder §41 die Weiterführung des Haushalts nicht möglich ist. Voraussetzung ist, dass im Haushalt ein Kind lebt, das das zwölfte Lebensjahr noch nicht vollendet hat oder das behindert und auf Hilfe angewiesen ist. Für Verwandte und Verschwägte bis zum 2. Grad werden keine Kosten erstattet (BMAS 2007, S. 177)

beschriebenen Erträge des Sozialkapitals generiert wurden, so dass direkt auf Versorgungsmöglichkeiten des formellen Sektors zurückgegriffen wird.

- Es wurden Investitionen getätigt und auch entsprechendes Sozialkapital produziert. Jedoch werden Personen aus technischen Gründen nicht genannt, da die Ansprechpartner zur Zeit nicht verfügbar oder in der Lage sind, im Falle der Pflegeabhängigkeit Hilfe zu leisten.
- Es wurden zwar Ressourcenpotentiale dieser Form aufgebaut, jedoch sind die Beziehungen zu diesen Personen derzeit vernachlässigt, gestört oder aber bereits nicht mehr vorhanden.
- Es existiert eine Präferenz des Befragten für professionelle Dienste im Vergleich zu den Unterstützungspotentialen der informellen Beziehungen.

Welches die korrekte Interpretation in der Situation des Befragten ist, kann nicht hergeleitet werden. Dies lässt sich exemplarisch an der Nennung der ersten Person darstellen, bei der sowohl die Möglichkeit bestand, „Niemand“ als auch ein „Substitut“ anzugeben. Beispielsweise haben 601 Personen ein Substitut in der ersten Nennung angegeben und bei den anderen beiden Antworten keine weiteren Angaben gemacht. Eine Präferenz für eine Marktlösung scheint hier eher abwegig; diese Antwort kann auch dahingehend interpretiert werden, dass bei diesen Personen tatsächlich keine Ansprechpartner vorhanden sind.

Gleichermaßen muss eine Person, die keine Ansprechpartner, hier also „Niemand“ nennt, zur Lösung der Problemlage pflegerische Versorgung, die normalerweise nicht alleine bewältigt werden kann, zwangsläufig auf ein Substitut zurückgreifen. Es bleibt unklar, welches die Beweggründe der Befragten für die unterschiedlichen Angaben sind. Zwischen diesen beiden Antworten kann inhaltlich nicht differenziert werden.

1. Analyse der Rangordnungen

Nach der Analyse der einzelnen Nennungen werden in diesem Abschnitt die Rangordnungen der Nennungen untersucht. Die Folge der genannten Ansprechpersonen aus Sicht der Befragten beschreibt, an wen sie im Falle einer Pflegebedürftigkeit als erstes, an wen als zweites und an wen sie als drittes denken. Durch die geäußerte Präferenz wird von den Befragten zweierlei zum Ausdruck gebracht. Erstens wird eine persönliche Rangordnung der Ansprechpersonen im Falle eines pflegerischen Bedarfs erzeugt. Zweitens wird angezeigt, welche Investitionsstrategien getätigt wurden, denn es ist davon auszugehen, dass zum Zeitpunkt der Befragung der Auskunft gebende davon überzeugt ist, dass diese Person auch als Ansprechpartner dient, da entsprechende Vor-

leistungen getätigt wurden und eine Verpflichtung zu einer derartigen Gegenleistung entstanden ist.

3.4 Datenprobleme

Die Frage des SOEP bezieht sich nur auf die drei ersten Ränge der sechs Personen-
gruppen, die wir im letzten Abschnitt unterschieden haben. Zudem haben nicht alle Per-
sonen eine komplette Rangfolge angegeben, sondern viele nur einen oder zwei An-
sprechpartner genannt. Warum die befragten Personen ihre Angaben nicht vervoll-
ständigt haben, kann nicht nachvollzogen, sondern es können bestenfalls Vermutungen
angestellt werden. Über die Hintergründe dieses Antwortverhaltens liegen keine hin-
reichenden Informationen vor. In Tabelle 4-1 werden aus Gründen der Übersichtlichkeit
ausgewählte Rangfolgen der Angaben dargestellt, wie sie von ca. zwei Drittel der Be-
fragten angegeben wurde. Eine vollständige Übersicht aller Antwortkonstellationen fin-
det sich in Tabelle 7-1 im Anhang. Jedes Analyseverfahren muss versuchen, auch diese
unvollständigen Rangangaben angemessen zu berücksichtigen.

Tabelle 4-1: Rangfolgen der Nennungen (ausgewählte Ergebnisse)

lfd. Nr.	Rangfolge der Nennung	Häufigkeiten	Prozent
1	1..	3339	23,88
2	12.	1663	11,89
3	122	1299	9,29
4	124	461	3,30
5	17.	722	5,16
6	2..	714	5,11
7	7..	601	4,30
Summe		6301	62,93

Quelle: SOEP 2006; Legende: 1 „(Ehe-)Partner“; 2 „Kind“; 3 „Eltern“; 4 „Verwandte“; 5 „Freunde, Nach-
barn, Sonstige“; 7 „Substitut“; . „fehlende Angabe“.

Aus analytischen Gesichtspunkten sind zudem gewisse, sich an dieser Stelle mani-
festierende Inkonsistenzen in den Antworten problematisch, da einige Personen identi-
sche Nennungen über zwei oder drei der Antworten gemacht haben. Im Sinne der Fra-
gestellung sollten solche Indifferenzen eigentlich ausgeschlossen werden. Beispiels-
weise haben 26 der Befragten den Partner mindestens zweimal als Ansprechperson ge-
nannt, obwohl sie, wenn man diese Angaben im Datensatz kontrolliert, überwiegend in
einer festen Partnerschaft im gemeinsamen Haushalt leben und gleichzeitig angeben,
keine ehemaligen Partner zu haben. Die Antworten beziehen sich offenbar auf ein und

dieselbe Person. Angegeben werden sollten aber bis zu drei unterschiedliche Personen, die bedeutsam für die genannte Fragestellung sind.

Ähnliche Angaben finden sich auch bei den anderen Antwortmöglichkeiten, z. B. bei der Nennung der Substitute, die zwei- oder dreifach präferiert wurden. Existiert in jedem Fall eine außerordentliche Präferenz für eine Marktlösung zur pflegerischen Versorgung, oder können hier verschiedene Formen der Substitution gemeint sein? Es kann sich aber auch um Personen handeln, deren Beweggründe zur Antwort identisch sind mit denen, die niemanden angegeben haben. Auch diese Formen redundanter Antworten können nicht in die Analyse der Rangfolgen einfließen, da sie mehrere Deutungen zulassen.

Unproblematisch sind solche Mehrfachnennungen theoretisch bei den Angaben zu Kindern, Eltern oder Verwandten als Ansprechpersonen unter der Voraussetzung, dass diese auch mehrfach vorhanden sind und kein Bezug auf dieselbe Person genommen wird. Prinzipiell könnte man die Werte insoweit korrigieren bzw. vervollständigen, indem bei identischen Nennungen z. B. bei den Kindern nach dem ersten, dem zweiten Kind, usw. unterschieden wird. Die dazu notwendigen Informationen über alle relevanten Personen sind jedoch nicht vollständig im Datensatz enthalten.

3.5 Statistische Analysen und explorative Verfahren

Hält man trotz dieser Einschränkungen an der Analyse fest, besteht die grundsätzliche Idee, zur statistischen Analyse solcher Rangfolgen den von den Befragten erzeugten Ordnungen numerische Werte, z.B. Rangzahlen zuzuordnen, die deren Angaben angemessen repräsentieren. Mittels einer solchen Zuordnung lassen sich Unterschiede und Abstände im mehrdimensionalen Merkmalsraum der zu Grunde liegenden Daten einer weiteren empirischen Analyse zugänglich machen, in der beispielsweise Subgruppen der Rangordnungen miteinander verglichen werden. Da auch Rangdaten multivariate Daten darstellen, lassen sie sich mit den dafür geeigneten Methoden analysieren (vgl. Marden 1995).

Dabei ist zu beachten, dass die vorhandenen Rangfolgen keine strikt linearen Ordnungen darstellen, da genannte Indifferenzen vorhanden sind. Dies schränkt die Aussagekraft der Rangordnung ein, da sich eine Größenrelation zwischen Ordnungen zwar prinzipiell erstellen lässt, eine darauf basierende vergleichende Wertung allerdings willkürlich ist. Es handelt sich also nicht um eine vollständige, sondern um eine partielle

Rangordnung, bei der nicht mehr alle Elemente im Sinne einer Ordnungsrelation miteinander verglichen werden können (vgl. umfassend: Rohwer et al. 2002). Eine vollständige Angabe über alle drei Nennungen haben 4662 Personen gemacht, über mindestens die ersten beiden Ansprechpartner haben 8904 der Befragten eine Auskunft gegeben (vgl. Tabelle 7-1 im Anhang).

Es bietet sich in einem ersten Schritt an, die Rangordnungen graphisch darzustellen, um einen Einblick in die Häufigkeitsverteilungen und die komplexe Struktur der Rangordnungen zu bekommen. Die Visualisierung mit Hilfe eines Polytops bietet eine geeignete Methode zur Darstellung der Mehrdimensionalität der Variablenausprägungen, auf deren Basis Erkenntnisse über adäquate Verfahren zur weiteren Analyse gewonnen werden können (vgl. Thompson 1993). Dazu kann man die Rang- oder Ordnungsvektoren als Vektoren im Euklidischen Raum auffassen. Für drei Alternativen ergibt sich etwa ein Bild wie in der folgenden Abbildung:

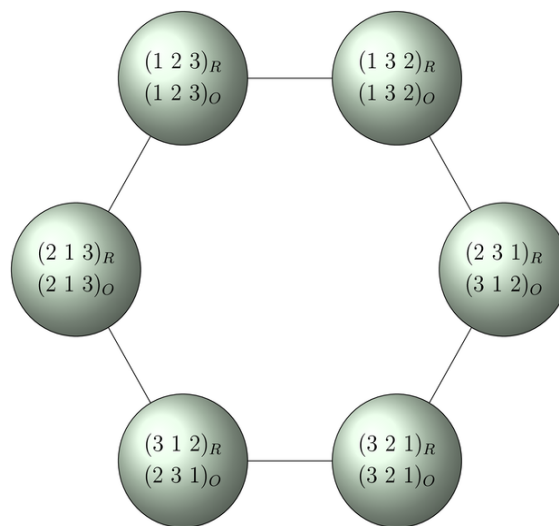


Illustration 1: Permutationspolytop für drei Alternativen.

Man erhält ein Polygon mit sechs Ecken, die wir sowohl mit den zugehörigen Ordnungs- und Rangvektoren bezeichnen. Dieses Polygon liegt auf der Ebene $x + y + z = 6$ im dreidimensionalen Raum. Je zwei benachbarte Punkte unterscheiden sich durch eine Vertauschung benachbarter Positionen der Ordnungsvektoren: Der Vektor $(123)_O$ und der rechts davon gelegene Vektor $(132)_O$ unterscheiden sich in den letzten beiden Stellen, der links von $(123)_O$ liegende Vektor $(213)_O$ in den ersten beiden Stellen. Den Abstand zwischen verschiedenen Ordnungen kann man nun als die kleinste Anzahl von Kanten definieren, die zwei Ordnungen verbindet. Diese Abstandsdefinition wird *Kendalls τ* genannt.

Was ergibt sich bei vier Alternativen? Das Permutationspolytop ist nun ein dreidimensionales Polytop im vierdimensionalen Raum, es wird von den Ordnungsvektoren $(1234)_O, (1243)_O, \dots$ aufgespannt und liegt in der dreidimensionalen Hyperebene $x+y+z+u=10$. Geometrisch ist es ein Oktaeder, dessen Ecken abgeschnitten wurden (Illustration 2):

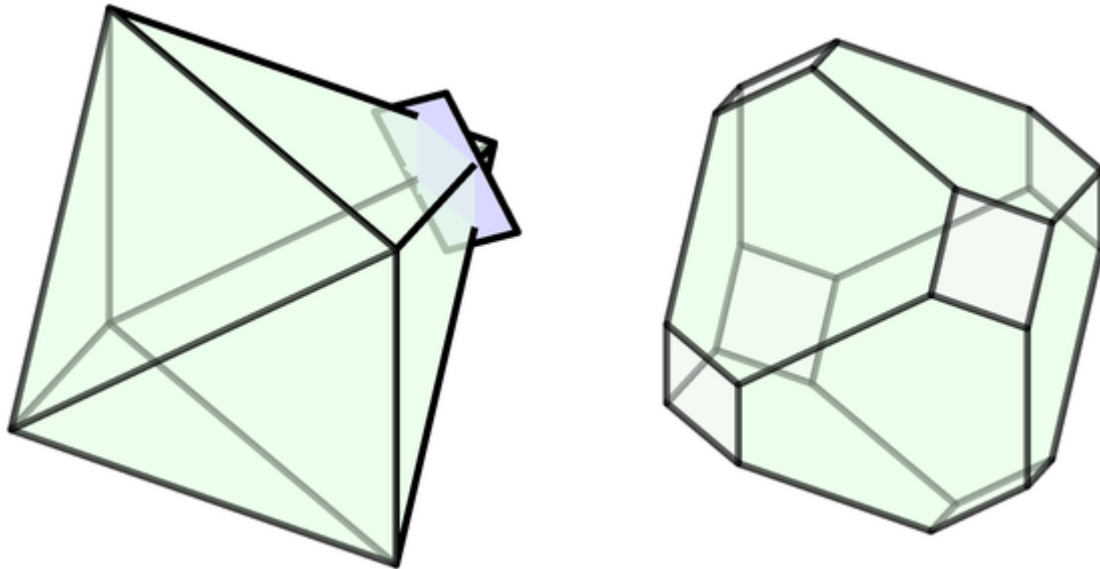


Illustration 2: Konstruktion des Permutationspolytop für vier Alternativen: Einem Oktaeder werden die Ecken abgeschnitten.

Belegt man die Ecken des Permutationspolytops mit den entsprechenden Ordnungsvektoren, dann erkennt man wieder, dass zwei Ecken genau dann durch eine Kante verbunden sind, wenn sich die zugehörigen Ordnungsvektoren nur durch eine Vertauschung benachbarter Stellen unterscheiden (vgl. Illustration 3). Entsprechend kann man Kendalls τ Abstand zwischen zwei Ordnungsvektoren wieder direkt ablesen.

Einen guten Überblick über die Verteilung von Rangdaten gewinnt man nun, wenn man an die Ecken des Permutationspolytops die Häufigkeit der entsprechenden Wahlen anträgt, etwa indem man Kugeln entsprechender Größe an die Ecken des Polytops malt. Das werden wir für die Rangdaten des SOEP in den nächsten beiden Abschnitten tun. Dazu müssen aber noch zwei offensichtliche Probleme gelöst werden: Wir unterscheiden in unseren Rangdaten sechs Alternativen. Das zugehörige Permutationspolytop liegt also im fünfdimensionalen Raum und lässt sich nicht mehr einfach zeichnen.⁵ Zudem haben wir nur unvollständige Rangdaten: Höchstens die ersten drei Präferenzen der

⁵ Zwar lassen sich auch noch vierdimensionale Polytope mit Schlegeldiagrammen halbwegs übersichtlich darstellen (vgl. Ziegler, 1995, Kap. 5), bei höheren Dimensionen muss man aber nach Alternativen suchen.

Befragten sind bekannt.

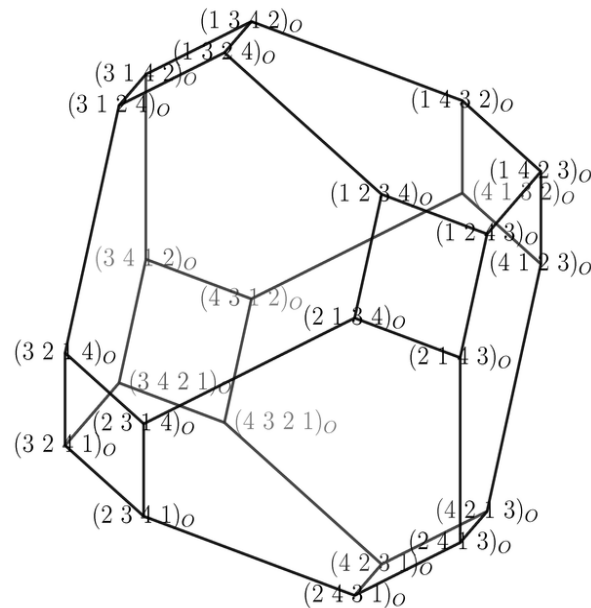


Illustration 3: Permutationspolytop für vier Alternativen und entsprechend gekennzeichneten Ecken.

Das erste Problem wird oft dadurch behoben, dass verschiedene Projektionen in den zwei- bzw. dreidimensionalen Raum betrachtet werden. Man kann etwa multidimensionale Skalierung oder Clusterverfahren verwenden. Der Nachteil solcher Verfahren ist, dass dabei die räumliche Struktur der Permutationspolytope mit Eigenschaften der Verteilung der Ränge vermischt werden. Andere Projektionsmethoden benutzen die Untergruppen der Permutationsgruppe. Dann aber bleibt nur ein Teil der Struktur erhalten. Unser Ziel hier ist aber die möglichst informative Darstellung der Häufigkeiten von Präferenzordnungen. Insbesondere sollen auch Abstände zwischen den Präferenzordnungen (zumindest Kendalls τ) erkennbar bleiben. Dazu reduzieren wir das Polytop auf seinen Graphen, d.h. wir behalten nur den Zusammenhang von Ecken und Kanten, wie es schon Thompson (1993) vorgeschlagen hat.⁶ Damit bleibt die Information über die Abstände und die Kantenzusammenhänge zwischen den Ecken erhalten. Damit bleibt auch die Information über Kendalls τ (und ähnlich konstruierte Metriken wie Cayleys Metrik, vgl. Marden, 1995, Kap. 2.5) erhalten. Zudem kann die Transformation unabhängig von den jeweiligen Daten sinnvoll durchgeführt werden. Auf der anderen Seite geht die Information über (höherdimensionale) Seiten und deren Zusammenhang ebenso wie die Information über Euklidische Abstände zwischen den Ecken, die u.a. für

⁶ Der Kantengraph eines Polytops wird auch manchmal das (eindimensionale) Gerüst des Polytops genannt.

die Berechnung von Spearmans ρ verwendet werden, verloren. Nun ist der Graph eines Polytops ab Dimension vier nicht mehr planar, lässt sich also nicht mehr überschneidungsfrei in der Ebene zeichnen. Dagegen kann er immer im dreidimensionalen Raum dargestellt werden, sogar so, dass alle Kanten durch Geradenstücke dargestellt werden können (Thomas, 2006, Kap. 4). An den Ecken dieses dreidimensional gezeichneten Graphen kann nun wieder die Häufigkeit der gewählten Präferenzordnungen angetragen werden.

Das zweite Problem der nur partiellen Ränge kann ebenso mit verschiedenen Methoden angegangen werden. Einen einheitlichen Gesichtspunkt gewinnt man, wenn man annimmt, dass es zu jeder partiellen Rangordnung eine wohl definierte Teilmenge aller vollständigen Ordnungen (bzw. aller Ränge) gehört, so dass die Elemente der Teilmenge mit der unvollständigen Präferenzordnung kompatibel sind.⁷ Z.B. wäre eine Angabe nur der ersten Präferenz aus (Partner, Verwandte, Freunde), etwa (Partner) kompatibel mit den Ordnungsvektoren $(1,2,3)_O$ und $(1,3,2)_O$. Es ist dann auch nahelegend, wie die geometrische Übersetzung aussehen sollte: Man nimmt den Durchschnitt (den Schwerpunkt) dieser kompatiblen Vektoren als neuen Punkt. Die aus den so definierten Eckpunkten erzeugten Polytope haben die gleiche (affine) Dimension wie die Polytope, die zu den entsprechenden vollständigen Präferenzordnungen gehören.⁸

Allerdings ist es bei partiellen Präferenzordnungen einfacher und konzeptionell klarer, von Rangvektoren auszugehen. Dann kann nämlich den Alternativen, deren Rang nicht angegeben wurde, ein einheitlicher Pseudorang zugewiesen werden. Bei den Ordnungsvektoren würde man dagegen verschiedenen Alternativen jeweils neue Namen zuordnen müssen. Gibt es etwa 6 Alternativen, von denen die ersten drei Präferenzen angegeben werden, dann ist der durchschnittliche Ordnungsvektor, der die Alternativen 1, 4 und 5 an die ersten drei Positionen setzt, $(1,4,5,11/3,11/3,11/3)_O$. Stehen dagegen die Alternativen 1, 2 und 3 an den ersten Positionen, dann erhält man $(1,2,3,5,5,5)_O$. Für

7 Marden (1995, Kap. 11) geht noch einen Schritt weiter und untersucht alle Teilmengen der Potenzmenge der Ränge als „unvollständige“ Rangdaten. Zhang (2004) benutzt das Permutationspolytop und unvollständige Präferenzordnungen verschiedener Art, um Zusammenhänge zwischen Präferenzordnungen, Auswahl von Teilmengen und verschiedenen Wahlverfahren zu beschreiben.

8 Daher können im Prinzip auch Daten, die sowohl vollständige als auch partielle Präferenzordnungen enthalten, graphisch dargestellt werden. Allerdings erhält man dann keine neuen Polytope: Alle Polytope partieller Präferenzordnungen sind echte Teilmengen des Permutationspolytops. Daher ist der Graph des Permutationspolytops nicht mehr informativ für die partiellen Präferenzen.

die Rangvektoren erhält man im ersten Fall $(1,5,5,2,3,5)_R$ und im zweiten Fall $(1,2,3,5,5,5)_R$.

Die minimale Anzahl der Kanten, die zwei Eckpunkte verbinden, ist wieder eine Metrik auf den partiellen Präferenzordnungen. Sie ist eine untere Schranke für Kendalls τ zwischen den kompatiblen vollständigen Präferenzen.⁹ Diese Metrik ist zudem invariant, wenn Alternativen umbenannt werden. Sie wird in den folgenden Anwendungen benutzt.

3.6 Die Rangordnungen der ersten beiden Ansprechpartner

Eine Differenzierung zwischen Personen, die nur die höchsten zwei Präferenzen angeben, und denjenigen, die die gewünschten drei wichtigsten Präferenzen äussern, empfiehlt sich aufgrund der anteilmäßig hohen Ausfälle in der Beantwortung. Vorab wird darauf hingewiesen, dass die Darstellung einer Problembeschreibung der vorliegenden Rangdaten dient. Zur Veranschaulichung dieser Problemlage werden daher die Mehrfachnennungen ausgeschlossen. Weiterhin wird vereinfachend angenommen, dass prinzipiell alle Personen in der Lage sind, Ansprechpartner aus den gegebenen Antwortmöglichkeiten zu nennen.

Die Antwortmöglichkeiten der ersten beiden Ränge bilden eine partielle Rangordnung. In diesem Fall gilt es, zwei Personen aus sechs Möglichkeiten auszuwählen. Unter den genannten Bedingungen erhält man für diese Permutation theoretisch dreißig verschiedene Antwortmöglichkeiten mit unterschiedlichen Kombinationen der ersten und zweiten Ansprechperson, die auch praktisch alle in den Daten vertreten sind (vgl. Tabelle 4-2).

In Abbildung 4-1 sind diese Permutationen der ersten und zweiten Nennung anhand eines Polytops dargestellt.¹⁰ Die Permutationen sind jeweils den Scheitelpunkten der entsprechenden Kanten zugeordnet. Identische Nennungen, also Bindungen¹¹, bei der ersten Person, bilden gemeinsam einen Pol. Inverse Nennungen, also diskordante Paare, besetzen gemeinsam die Endpunkte einer Kante. Durch die Größe der Kugeln an den Ecken werden die Häufigkeiten der Nennungen repräsentiert.¹²

Nicht zu übersehen ist die überragende Bedeutung der Partner-Kind Nennung (PK) in

⁹ Kidwell et al. (2008) haben vorgeschlagen, den durchschnittlichen Wert von Kendalls τ als Metrik zu definieren. Sie diskutieren auch kurz weitere Metriken für partiellen Präferenzangaben. Thompson (1993) diskutiert die hier vorgestellte Variante.

den ersten beiden Rangordnungen. Diese ist so eminent, dass von dieser Antwortkombination der gesamte Pol vereinnahmt wird. Es muss nochmals betont werden, dass in dieser Darstellung von einer Beschränkung auf die Personen, die auch über alle Ansprechpartner verfügen, abgesehen wurde. Würde man auf den Personenkreis eingrenzen, der mindestens über Partner und Kind verfügt, wäre die hier gewählte grafische Darstellung überhaupt nicht mehr möglich, da diese beiden Nennungen mit Abstand am häufigsten erfolgt ist.

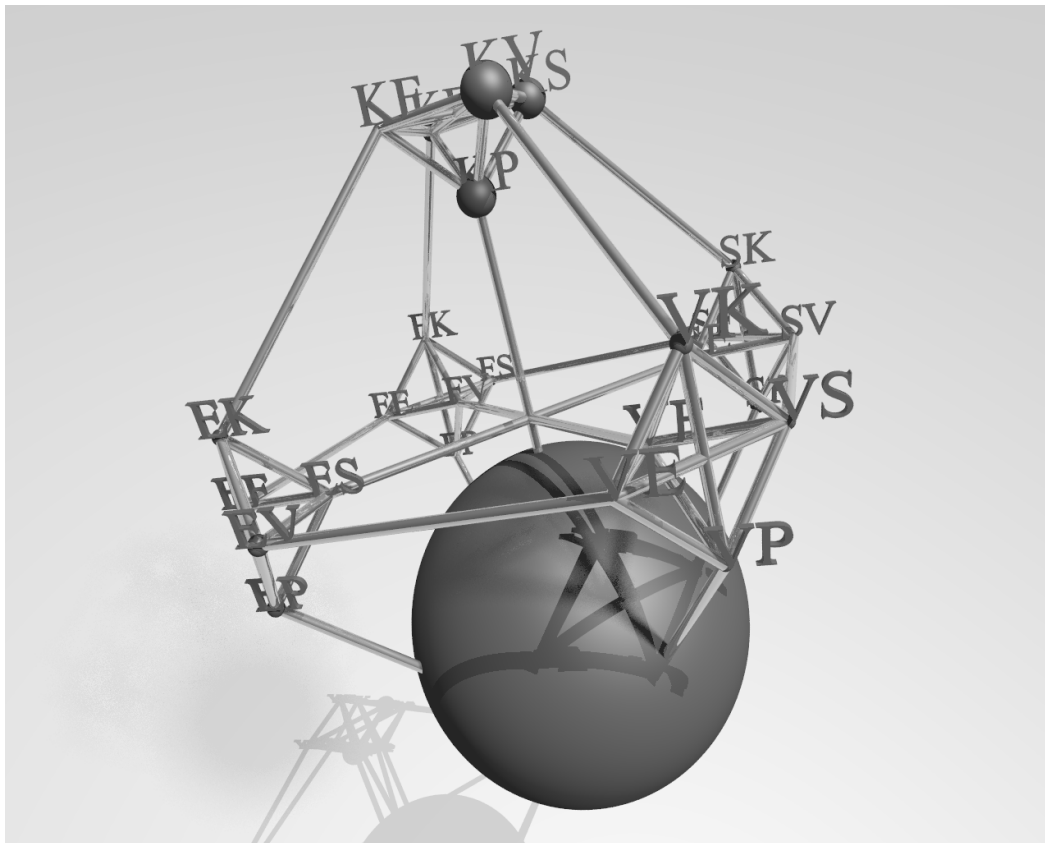
Kumuliert man die Häufigkeiten der Rangfolgen mit einem Partner an der ersten Stelle, so erfolgen an diesem Pol über 80 Prozent der Nennungen (Tabelle 4-2). Im Vergleich zu den übrigen Rangfolgen erfahren noch die Ränge mit Kindern als erste Ansprechperson eine gewisse Bedeutung, die aber anteilsmäßig lediglich zwischen zwei und drei Prozent liegen.

10 Die technische Umsetzung der Abbildungen erfolgte auf folgendem Weg: Die Permutationen wurden mit Hilfe des Statistikprogramms R (Version 2.7.0; <http://www.r-project.org/>) in Form einer Matrix abgebildet. Die Koordinaten der Eckpunkte und Kanten für die Polytope wurden mit der Software Polymake (Version 2.3; <http://www.math.tu-berlin.de/polymake/>) berechnet. Für die Weiterbearbeitung und Visualisierung dieser Werte diente das Programm Povray (Version 3.6.1; <http://www.povray.org/>). Die gemeinsame Nutzung der Programme basierte auf dem Betriebssystem Linux (OpenSuse 10.3; <http://de.opensuse.org/>). Alle Programme sind unter den angegebenen Adressen frei zugänglich.

11 In der Literatur werden neben dem Begriff der Bindungen diese gleichermaßen als Verknüpfungen oder englisch als „ties“ bezeichnet (vgl. z. B. Benninghaus 1992; Rohwer et al. 2002).

12 Man beachte in der Grafik, dass die Darstellung der Häufigkeiten eine Frage der Skalierung in der Programmierung des Polytops ist und die Radien der Kugeln unterschiedlich dimensioniert werden können. Im vorliegenden Fall müsste gemäß der Häufigkeiten (vgl. Tabelle 4-2) die Kugel PK (Partner-Kind) größer sein, was der Visualisierung des Problems allerdings abträglich ist. Aus Gründen einer verbesserten Darstellung wurden die Radien entsprechend anders skaliert, die Proportionen der Kugeln untereinander bleiben davon unberührt. Existieren an den Ecken in der Grafik keine Kugeln, bedeutet dies lediglich, dass sich aufgrund geringer Häufigkeiten keine sichtbare Kugel bildet.

Abbildung 4-1: Die Rangordnungen der ersten beiden Ansprechpersonen



Quelle: SOEP 2006; Legende: „E“ Eltern, „P“ Partner, „K“ Kinder, „V“ Verwandte, „S“ Substitut, „F“ Freunde.

Tabelle 4-2: Absolute und relative Häufigkeiten der ersten beiden Ränge

Rangfolge	EF	EK	EP	ES	EV	FE	FK	FP	FS	FV
n	26	43	54	26	58	8	24	12	41	21
%	0,32	0,53	0,67	0,32	0,72	0,1	0,3	0,15	0,51	0,26
Rangfolge	KE	KF	KP	KS	KV	PE	PF	PK	PS	PV
n	21	64	180	164	223	1002	252	4048	850	602
%	0,26	0,79	2,22	2,03	2,76	12,38	3,11	50,03	10,51	7,44
Rangfolge	SE	SF	SK	SP	SV	VE	VF	VK	VP	VS
n	7	53	58	65	16	13	42	53	27	38
%	0,09	0,66	0,72	0,8	0,2	0,16	0,52	0,66	0,33	0,47
n gesamt	8091									

Quelle: SOEP 2006; Legende: E „Eltern“, P „Partner“, K „Kinder“, V „Verwandte“, S „Substitut“, F „Freunde“.

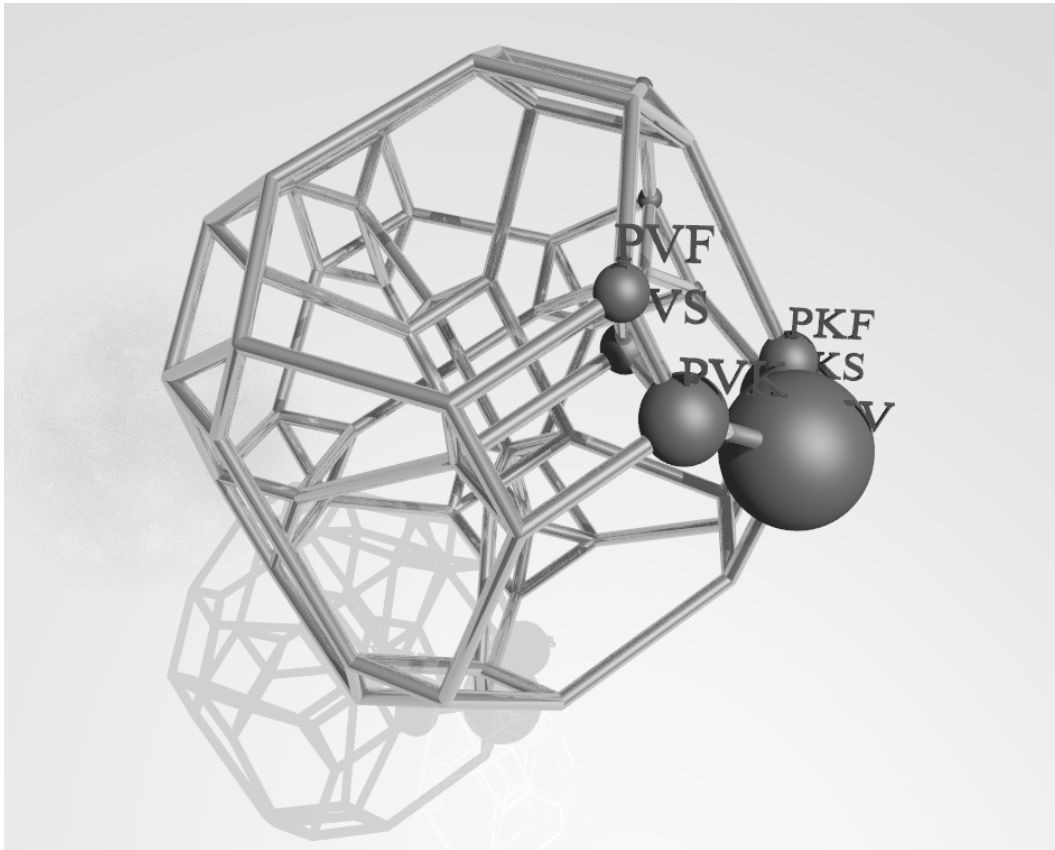
Das offensichtliche Resultat lässt sich knapp zusammenfassen: Die von den Befragten angegebene Rangordnung ist bei der Hälfte der Befragten identisch, bei weiteren 33 Prozent ist bei demselben ersten Rang nur der zweite Rang verschieden. Es verbleiben 17 Prozent der Befragten, die eine gänzlich alternative Präferenz geäußert haben.

3.7 Analyse der vollständigen Rangfolge

Die vollständige Rangfolge lässt sich ebenfalls mit dem bereits beschriebenen Verfahren visualisieren. In Unterschied zur ersten Darstellung werden nun die Eltern den Verwandten zugeordnet. Die daraus resultierende Permutation drei Personen aus fünf möglichen Personengruppen zu wählen, bietet 60 Auswahlmöglichkeiten, wenn man wiederum identische Nennungen ausschließt. Eine solche Angabe haben 2040 der befragten Personen gemacht.

Bis auf die Kombination „Freunde-Partner-Substitut“ (FPS; vgl. Tabelle 4-3) wurden alle Rangfolgen mindestens einmal angegeben. Allerdings kommt außerhalb der Rangordnungen mit dem Partner in der ersten Besetzung den weiteren Möglichkeiten eine eher marginale Bedeutung zu¹³, was letztlich in der Grafik dazu führt, dass sich an den entsprechenden Stellen im Polytop keine sichtbaren Kugeln formieren. Der Schwerpunkt der Nennungen liegt im „Partnerpol“, der in seiner Gesamtheit auch hier mit ca. 80 Prozent die überragende Mehrheit der Befragten repräsentiert. Personen, die eine vollständige Rangfolge angegeben haben, scheinen überwiegend in partnerschaftlichen Beziehungen eingebunden zu sein und Kinder zu haben. Man könnte nun formulieren, dass dies der wesentliche Grund ist, dass eine vollständige Rangfolge angegeben, weil familiär eingebundene Personen beziehungs technisch umfassend versorgt sind. Aufgrund der geschilderten Problematik auf der inhaltlichen Ebene im Rahmen der Fragestellung scheint ein solcher Schluss aber gewagt.

¹³ Aufgrund der geringen Zahl der Nennungen wurde in der Grafik auf die Darstellung der Labels bei diesen Antwortmöglichkeiten verzichtet.

Abbildung 4-2: Die Rangordnungen über alle Ansprechpersonen

Quelle: SOEP 2006; Legende: P „Partner, K „Kinder“, V „Verwandte, S „Substitut, F „Freunde“.

Tabelle 4-3: Absolute und relative Häufigkeiten der Ränge

Rangfolge	FKP	FKS	FKV	FPK	FPS	FPV	FSK	FSP	FSV	FKV
n	1	3	1	1	0	3	2	2	3	3
%	0.05	0.15	0.05	0.05	0	0.15	0.1	0.1	0.15	0.15
Rangfolge	FVP	FVS	KFP	KFS	KFV	KPF	KPS	KPV	KSF	KSP
n	1	2	1	11	3	4	19	19	19	4
%	0.05	0.1	0.05	0.54	0.15	0.2	0.93	0.93	0.93	0.2
Rangfolge	KSV	KVF	KVP	KVS	PKF	PFS	PFV	PKF	PKS	PKV
n	5	18	16	26	12	40	16	158	344	581
%	0.25	0.88	0.78	1.27	0.59	1.96	0.78	7.75	16.86	28.48
Rangfolge	PSF	PSK	PSV	PVF	PVK	PVS	SFK	SFP	SFV	SKF
n	59	44	18	119	233	113	6	1	6	4
%	2.89	2.16	0.88	5.83	11.42	5.54	0.29	0.05	0.29	0.2
Rangfolge	SKP	SKV	SPF	SPK	SPV	SVF	SVK	SVP	VFK	VFP
n	5	6	6	13	6	1	2	0	1	2
%	0.25	0.29	0.29	0.64	0.29	0.05	0.1	0	0.05	0.1
Rangfolge	VFS	VKF	VKP	VKS	VPF	VPK	VPS	VSF	VSK	VSP
n	13	3	12	9	6	13	6	10	2	3
%	0.64	0.15	0.59	0.44	0.29	0.64	0.29	0.49	0.1	0.15
n gesamt	2040									

Quelle: SOEP 2006; Legende: „P“ Partner, „K“ Kinder, „V“ Verwandte, „S“ Substitut, „F“ „Freunde“.

Die geschilderte Problematik ist offensichtlich und braucht nicht viel weiterer Erklä-

nung: Die hohen Ausfälle in den Daten, d.h. Personen, die eine unvollständige Rangfolge abgegeben haben, ist hochgradig selektiv. Es verbleiben fast ausschließlich Personen für die Analyse der vollständigen Rangordnungen, die über einen Partner und Kinder verfügen, und die Hilfestellung dieser auch präferieren. D. h., dass eine idealerweise vorhandene Gleichverteilung, bei der alle Rankings eine gleiche Wahrscheinlichkeit besitzen, nicht erfüllt ist (vgl. Marden 1995, Kapitel 3). Auch unter Hinzunahme der Mehrfachnennungen kann dieses Problem nicht gelöst werden, die Selektivität bleibt erhalten (vgl. Tabelle 7-1). Unterschiede in den Investitionsstrategien in Sozialkapital lassen sich damit empirisch nicht mehr fassen.

4 Schlussbetrachtung und Ausblick

Die Analyse der hier verwendeten Daten hat gezeigt, dass bei Verwendung einer Fragekonstellation auf der Erwartungsebene des Sozialkapitals unter Umständen ein inhaltlich nicht eindeutig zu interpretierendes Ergebnis produziert wird, was die Analysemöglichkeiten erheblich einschränkt. Dabei wurde deutlich, welche Ambivalenz sich hinter einem Erwartungswert verbirgt. Der empirische Erkenntnisgewinn wird durch eine stark eingeschränkte Aussagekraft gemindert, in der die supportiven Aspekte des Sozialkapitals nicht mehr eindeutig ermittelt werden können.

Zusammenfassend lässt sich somit feststellen, dass die offen formulierte Fragestellung, in der nach drei Ansprechpersonen im Falle einer hypothetischen Pflegebedürftigkeit gefragt wurde, erstens ein Ergebnis produziert, dass aus Sicht des Forschers inhaltlich nicht zweifelsfrei zu identifizieren ist. Zweitens provoziert die Ungenauigkeit der Fragestellung bei den Befragten ein Antwortverhalten, das letztlich zu einer Konzentration der geäußerten Präferenzen auf wenige Antworten führt, die mit den zur Verfügung stehenden statistischen Methoden nicht mehr sinnvoll untersucht werden können. Diese inhaltlichen und methodischen Probleme führen aus den genannten Gründen dazu, dass an dieser Stelle von einer weiteren Analyse abgesehen werden muss.

Ein weiteres Problem, dass hier ausführlich behandelt wurde, stellen teilweise unvollständige, aber aufeinander zu beziehende Angaben der Befragten dar. Mit den konventionellen Verfahren der Statistik können diese Daten nicht ohne weitere Bemühungen untersucht werden. Anhand der Darstellung von Rangordnungen in Form eines Polytops konnte explorativ veranschaulicht werden, dass Analysemöglichkeiten durchaus vorhanden sind, die ausbaufähig sind. Mit diesen, bislang in der sozialwissenschaftlichen

Forschung wenig beachteten Verfahren lassen sich (partielle) Rangordnungen, wie sie in Datensätzen nicht selten zu finden sind, der empirischen Analyse und weiterführenden Untersuchungen zugänglich machen. Die besondere Stärke dieser visuellen Darstellung liegt insbesondere darin begründet, dass sich Strukturen in den Präferenzbeurteilungen prüfen und identifizieren lassen, was als Grundlage weiterer Überlegungen dienlich sein kann. Gleichzeitig können mit der Auswahl einer geeigneten Metrik anschließende multivariate Analysemöglichkeiten eröffnet werden.

Mit Blick auf zukünftige Befragungen und empirische Untersuchungen ist eine Diskussion über ihre Art und die Form der Erhebungsinstrumente wünschenswert. Wenn weiterhin an der Messung dieser Formen des Sozialkapitals festgehalten werden soll, muss thematisiert werden, inwieweit speziell Erwartungswerte sinnvolle Ergebnisse liefern können. Eine Verbesserung kann einerseits in der Modifikation der Erhebungsinstrumente liegen. Andererseits kann konkret messbares Sozialkapital, beispielsweise in Form aktivierter Ressourcenpotentiale eine Alternative oder aber sinnvolle Ergänzung darstellen. Mit einer solchen Vorgehensweise könnte auch den sich hier als problematisch herausgestellten hohen Ausfällen in den Fallzahlen begegnet werden, durch die die Verwertbarkeit der Daten erheblich eingeschränkt ist. Dem kann auch begegnet werden, wenn in der Zukunft den Befragten durch weiterreichende Antwortkategorien die Möglichkeit gegeben wird, ihre Lebenssituation genauer darzustellen, so dass die Gründe des Antwortverhaltens besser nachvollzogen werden können.

5 Literaturverzeichnis

- Benninghaus, H. (1992): Deskriptive Statistik. Stuttgart.
- Bundesministerium für Arbeit und Soziales (BMAS) (2007): Übersicht über das Sozialrecht. 4. Auflage, Nürnberg.
- Cohen, S./Syme, L., S. (1984): Issues in the Study and Application of Social Support. In: Cohen, Sheldon/Syme, Leonard, S. (Hrsg.): Social Support and Health. Orlando, Florida, London, S. 3-22.
- Diewald, M./Lüdicke, J./Lang, F. R./Schupp, J. (2006): Familie und soziale Netzwerke. Ein revidiertes Erhebungskonzept für das Sozioökonomische Panel (SOEP) im Jahr 2006. In: DIW Berlin Research Notes 14.
- Diewald, M. (1991): Soziale Beziehungen: Verlust oder Liberalisierung? Soziale Unterstützung in informellen Netzwerken. WZB Berlin.
- Hank, K./Erlinghagen, M. (2008): Produktives Altern und informelle Arbeit. In: Erlinghagen, M./Hank, K. (Hrsg.): Produktives Altern und informelle Arbeit in modernen Gesellschaften. Theoretische Perspektiven und empirische Befunde. Wiesbaden, S. 9-24.
- Kidwell, P./Lebanon, G./Cleveland, W. (2008): Visualizing incomplete and partially ranked data. IEEE Transactions on Visualization and Computer Graphics. Heft 14, S. 1356-1363.
- Marden, J., I. (1995): Analyzing and Modeling Rank Data. Monographs on Statistics and Applied Probability. Band 64, London.
- Pihl, C. (2008): Sozialkapital aus investitionstheoretischer Sicht. Dissertation an der Ruhr-Universität Bochum, Fakultät für Sozialwissenschaft.
- Rohwer, G./Pötter, U. (2002): Methoden sozialwissenschaftlicher Datenkonstruktion. Weinheim u.a..
- Sattler, S./Diewald, M. (2009): Kosten und Nutzen der Sparsamkeit - Zur Erhebung sozialer Netzwerke und sozialer Unterstützungspotentiale in der Umfrageforschung. SOEPpapers 165. Berlin: DIW Berlin.
- Thomas, R.R. (2006): Lectures on Geometric Combinatorics. AMS, Providence.
- Thompson, G., L. (1993): Generalized permutation polytopes and exploratory graphical methods for ranked data. In: The Annals of Statistics, Heft 21, Nr. 3, S. 1401-1430.
- Wagner, G.; Göbel, J.; Krause, P.; Pischner, R. und Sieber, I. (2008): Das Sozio-oekonomische Panel (SOEP): Multidisziplinäres Haushaltspanel und Kohortenstudie für Deutschland - Eine Einführung (für neue Datennutzer) mit einem Ausblick (für erfahrene Anwender). In: AStA Wirtschafts- und Sozialstatistisches Archiv, Jg. 2, Nr. 4, S. 301-328.
- Zhang, J. (2004): Binary choice, subset choice, random utility, and ranking: A unified

perspective using the permutahedron. Journal of Mathematical Psychology. Heft 48, S. 107-134.

Ziegler, G.M. (1995): Lectures on Polytopes, Berlin.

6 Anhang

Tabelle 7-1: Übersicht der Rangordnung der genannten Unterstützungspersonen bei einer längerfristigen Pflegebedürftigkeit

Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte
1..	3339	23,88	144	113	63,71	222	47	81,25	27.	131	85,53
11.	20	24,03	145	64	64,17	223	3	81,27	271	4	85,56
111	6	24,07	147	66	64,64	224	81	81,85	272	3	85,58
112	2	24,08	15.	145	65,67	225	30	82,06	274	5	85,62
113	10	24,15	152	12	65,76	227	46	82,39	275	19	85,75
114	5	24,19	153	8	65,82	23.	6	82,43	277	2	85,77
115	3	24,21	154	8	65,88	231	6	82,48	3..	105	86,52
12.	1663	36,11	155	39	66,15	232	1	82,48	31.	26	86,7
121	3	36,13	157	40	66,44	233	2	82,5	311	1	86,71
122	1299	45,42	17.	722	71,6	234	5	82,53	312	8	86,77
123	120	46,28	172	44	71,92	235	1	82,54	313	6	86,81
124	461	49,57	173	3	71,94	24.	113	83,35	314	5	86,85
125	158	50,7	174	15	72,05	241	10	83,42	315	4	86,88
127	344	53,17	175	59	72,47	242	8	83,48	317	4	86,9
13.	295	55,28	177	7	72,52	243	3	83,5	32.	18	87,03
131	1	55,28	2..	714	77,63	244	46	83,83	321	4	87,06
132	160	56,43	21.	118	78,47	245	17	83,95	322	7	87,11
133	287	58,48	212	20	78,61	247	26	84,14	323	1	87,12
134	157	59,6	213	5	78,65	25.	35	84,39	324	8	87,18
135	55	60	214	14	78,75	251	1	84,39	325	1	87,18
137	47	60,33	215	4	78,78	252	2	84,41	327	4	87,21
14.	244	62,08	217	19	78,91	254	3	84,43	33.	56	87,61
142	73	62,6	22.	245	80,67	255	12	84,51	331	12	87,7
143	42	62,9	221	34	80,91	257	11	84,59	332	13	87,79

Quelle: SOEP 2006, Legende: 1 „(Ehe-)Partner; 2 „Kind“; 3 „Eltern“; 4 „Verwandte“; 5 „Freunde, Nachbarn, Sonstige“; 7 „Substitut“; . „fehlende Angabe“; über 40-jährige Personen.

Fortsetzung von Tabelle 7-1:

Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte	Rangfolge der Nennung	Häufigkeiten	Kumulierte Prozentwerte
333	3	87,81	415	2	90,47	47.	28	92,26	55.	11	93,81
334	32	88,04	417	2	90,49	471	2	92,28	553	2	93,83
335	9	88,11	42.	21	90,64	472	1	92,28	554	2	93,84
337	9	88,17	421	8	90,69	474	1	92,29	555	13	93,93
34.	27	88,36	422	12	90,78	475	5	92,33	557	5	93,97
341	5	88,4	424	5	90,82	477	1	92,33	57.	28	94,17
342	2	88,41	425	2	90,83	5..	131	93,27	571	2	94,18
343	2	88,43	427	5	90,87	51.	8	93,33	572	2	94,2
344	11	88,51	43.	6	90,91	512	1	93,33	574	3	94,22
345	6	88,55	431	2	90,92	513	1	93,34	575	5	94,26
347	5	88,58	433	3	90,94	514	2	93,36	577	1	94,26
35.	12	88,67	435	1	90,95	52.	13	93,45	7..	601	98,56
351	1	88,68	437	1	90,96	521	1	93,46	71.	40	98,85
352	1	88,68	44.	57	91,37	522	5	93,49	712	13	98,94
355	6	88,73	441	5	91,4	524	1	93,5	713	4	98,97
357	6	88,77	442	7	91,45	525	1	93,51	714	2	98,98
37.	19	88,91	443	2	91,47	527	3	93,53	715	6	99,03
371	1	88,91	444	17	91,59	53.	4	93,56	72.	31	99,25
372	1	88,92	445	14	91,69	533	3	93,58	721	5	99,28
375	5	88,96	447	10	91,76	534	1	93,58	722	12	99,37
4..	187	90,29	45.	25	91,94	54.	13	93,68	723	1	99,38
41.	11	90,37	451	1	91,95	541	1	93,68	724	5	99,41
411	1	90,38	453	1	91,95	542	3	93,71	725	4	99,44
412	5	90,42	454	1	91,96	543	1	93,71	745	1	99,61
413	4	90,44	455	7	92,01	545	1	93,72	75.	35	99,86
414	2	90,46	457	7	92,06	547	2	93,73	751	1	99,86

Legende: 1 „(Ehe-)Partner; 2 „Kind“; 3 „Eltern“; 4 „Verwandte“; 5 „Freunde, Nachbarn, Sonstige“; 7 „Substitut“; . „fehlende Angabe“; über 40-jährige Personen.

Fortsetzung von Tabelle 7-1:

Rangfolge der Nen- nung	Häufig- keiten	Kumulierte Pro- zent-werte
752	6	99,91
753	2	99,92
754	4	99,95
755	5	99,99
77.	1	99,99
777	1	100
Total	13981	

Quelle: SOEP 2006; Legende: 1 „(Ehe-)Partner; 2 „Kind“; 3 „Eltern“; 4 „Verwandte“; 5 „Freunde, Nachbarn, Sonstige“; 7 „Substitut“; . „fehlende Angabe“; über 40-jährige Personen.